La media armónica. La media armónica es una medida que se utiliza preferentemente cuando los datos se encuentran en función del tiempo, es decir, sustituye a la media aritmética como indicador de posición cuando los datos se refieren a determinadas tasas de utilización. Esta medida se la denota como: M. H.

Calculo de la media armónica. Se divide en dos casos cuando los datos no están tabulados, y cuando los datos están tabulados.

Datos no agrupados. La media armónica es un valor de la variable que se obtiene como el reciproco de la media de la suma de los valores recíprocos de la variable, es decir:

$$M.H. = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

Ejemplo: Determine la media armónica de los datos: 4, 6, 10

$$M.H. = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3}} = \frac{3}{\frac{1}{4} + \frac{1}{6} + \frac{1}{10}} = 5.806$$

Datos agrupados. Se aplica una expresión similar con la diferencia de que se toma en cuenta las frecuencias correspondientes a cada una de las variables, es decir:

$$M.H. = \frac{n}{\sum_{i=1}^{m} \frac{n_i}{x_i}} \quad , m = numero\ de\ filas$$

Ejemplo: Determine la media armónica de

Xi	**n_i**	**h_i (%)**	**N_i**	**H_i**
50	4	11.43	4	11.43
51	4	11.43	8	22.86
52	12	34.28	20	57.14
53	8	22.86	28	80.00
54	7	20.00	35	100.00
TOTAL	**35**			

Solución

$$M.H. = \frac{n}{\sum_{i=1}^{m} \frac{n_i}{x_i}} = \frac{n}{\frac{n_1}{x_1} + \frac{n_2}{x_2} + \frac{n_3}{x_3} + \frac{n_4}{x_4} + \frac{n_5}{x_5}} = \frac{35}{\frac{4}{50} + \frac{4}{51} + \frac{12}{52} + \frac{8}{53} + \frac{7}{54}} = 52.256$$

Los cuartiles. Los cuartiles son valores que dividen a un conjunto de datos ordenados en forma ascendente o descendente en cuatro partes iguales, teniéndose de esta forma tres puntos de división para tener las cuatro partes, es en ese sentido que tendremos 3 cuartiles, es decir:

$$Q_1 = primer\ cuartil$$

$$Q_2 = segundo\ cuartil$$

$$Q_3 = tercer\ cuartil$$

Datos no agrupados. Para este caso

Primero para calcular los cuartiles ordenamos los valores de forma creciente, luego:

- Para el Q_1
  i) si $n = par$ entonces $Q_1 = X_{\left(\frac{n}{4}\right)}$
  ii) si $n = impar$ entonces $Q_1 = X_{\left(\frac{n+1}{4}\right)}$
- Para el segundo cuartil, el cual representa simplemente la mediana.

- Para el tercer cuartil Q_3
  i) si $n = par$ entonces $Q_3 = X_{\left(\frac{3*n}{4}\right)}$
  ii) si $n = impar$ entonces $Q_3 = X_{3*\left(\frac{n+1}{4}\right)}$
  Remarca: en caso de no incidir con una posición exacta se puede sacar el promedio entre los valores de las dos variables de posiciones más próximas o interpolar, en nuestro caso realizaremos el promedio

Ejemplo: Sean los datos

2, 5, 6, 10, 12, 14, 15, 17, 20, 21

Hallar Q_1 y Q_3

Como

$$n = 10 = par, Q_1 = X_{\left(\frac{n}{4}\right)} = X_{\left(\frac{10}{4}\right)} = X_{2.5} = \frac{5+6}{2} = 5.5$$

Como

$$n = 10 = par, Q_1 = X_{\left(\frac{3*n}{4}\right)} = X_{\left(\frac{30}{4}\right)} = X_{7.5} = \frac{15+17}{2} = 16$$

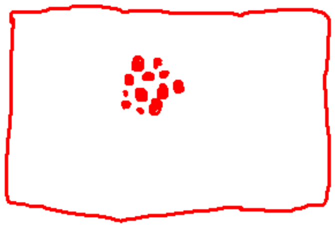
MEDIDAS DE DISPERSIÓN, COMPARACIÓN Y ASIMETRÍA

Medidas de Dispersión. Las medidas de dispersión son aquellas que permiten determinar el grado de concentración y el grado de dispersión de los valores de la variable respecto a una medida de posición.

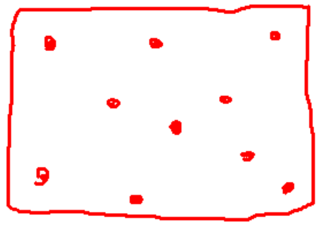
Medida de dispersión = ninguna = 0

Medida de dispersión = poca dispersión ≈ 0

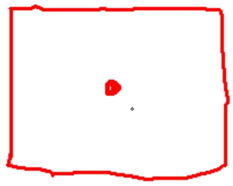
Medida de dispersión = mucha dispersión =número es grande



Poca dispersión (próximo a cero)



mucha dispersión (número grande)



Ninguna dispersión (cero)

Ejemplo: Sean los datos

4,4,4,4,4

Medida de dispersión =0

Ejemplo: Sean los datos

4, 5, 5.1, 5.2, 5.3

Medida de dispersión = próximo a cero

Las principales medidas de dispersión son:

- El recorrido de la variable.
- La desviación media absoluta.
- La varianza.
- La desviación típica o estándar

El recorrido de la variable. Es la medida más sencilla y de cálculo más fácil, la misma nos indica la dispersión global de los datos. Comúnmente se denota como: R o r

El cálculo del recorrido de la variable se efectúa mediante la diferencia del valor máximo y el valor mínimo de los valores observados por la variable, es decir:

$$R = r = x_{max} - x_{min}$$

Remarca: Este cálculo es para datos tabulados o no tabulados.

Ejemplo: Determine el recorrido de los datos

4, 5, 7, 10, 1, 29

$$R = r = x_{max} - x_{min} = 29 - 1 = 28$$

Ejemplo: Determine el recorrido de los datos

Xi	n_i	h_i (%)	N_i	H_i
5	4	11.43	4	11.43
7	4	11.43	8	22.86
8	12	34.28	20	57.14
13	8	22.86	28	80.00
27	7	20.00	35	100.00
TOTAL	**35**			

$$R = x_{max} - x_{min} = 27 - 5 = 22$$

Desviación media absoluta. Es la media aritmética en valor absoluto de las desviaciones de los valores observados de la variable respecto de la media aritmética. Se denota comúnmente como: DMA.

Se calcula mediante las expresiones:

- Para datos no tabulados

$$DMA = \sum_{i=1}^{n} \frac{|x_i - \mu|}{n},$$

- Para datos tabulados

$$DMA = \sum_{i=1}^{m} \frac{|x_i - \mu| * n_i}{n}, \quad m = numero\ de\ filas$$

La desviación media absoluta esta expresada en las mismas unidades de la variable y se puede utilizar como medida de dispersión en todas aquellas distribuciones en la que la medida de tendencia central más relevante es la media aritmética.

La desviación media presenta las siguientes ventajas:

- ✓ En su cálculo se utilizan todas las observaciones.
- ✓ Es menos sensible a la presencia de valores extremos respecto al recorrido de la variable.

La desviación media absoluta, presenta las siguientes desventajas:

- ✓ Su manejo algebraico es complicado.
- ✓ Su interpretación es relativamente difícil.

Ejemplo: Si tenemos los siguientes datos: 2, 4, 5, 8, 10, 10.
Determinar la desviación media absoluta

Sol: Para determinar la desviación media absoluta aplicaremos la siguiente expresión:

$$\mu = \frac{2 + 4 + 5 + 8 + 10 + 10}{6} = 6.5$$

$$MDA = \sum_{i=1}^{n} \frac{|x_i - \mu|}{n}$$

$$= \frac{|2-6.5| + |4-6.5| + |5-6.5| + |8-6.5| + |10-6.5| + |10-6.5|}{6} = 2.83$$

Ejemplo: Determine el Desviación media absoluta

[L_{i-1} - L_i)	**Xi**	**n_i**	**h_i (%)**	**N_i**	**H_i (%)**
[7 a 9.4)	8.2	8	32	8	32
[9.4 a 11.8)	10.6	5	20	13	52
[11.8 a 14.2)	13	4	16	17	68
[14.2 a 16.6)	15.4	6	24	23	92
[16.6 a 19]	17.8	2	8	25	100
TOTAL		25	100		

$$\mu = \frac{8.2*8 + 10.6*5 + 13*4 + 15.4*6 + 17.8*2}{25} = 11.944$$

$$DMA = \sum_{i=1}^{m} \frac{|x_i - \mu| * n_i}{n} = \frac{|x_1 - \mu| * n_1 + \cdots + |x_m - \mu| * n_m}{n}$$

$$= \frac{|8.2 - 11.944| * 8 + |10.6 - 11.944| * 5 + |13 - 11.944| * 4 + |15.4 - 11.944| * 6 + |17.8 - 11.944| * 2}{25}$$

$$= 2.934$$

Ejemplo: Sea los datos: 7,7,7,7

Media aritmética =7

D.M.A = 0

Varianza. La varianza se define como la media aritmética del cuadrado de las desviaciones de los valores de la variable respecto de la media aritmética.

La varianza se representa mediante: σ^2, $\sigma_x^2, V(X), VAR(X)$, S^2.

Calculo de la varianza. El cálculo de la varianza depende si los datos están tabulados o no, la varianza se halla mediante uno de los dos siguientes procesos:

Para datos no agrupados: Utilizaremos cualquiera de las dos siguientes expresiones.

$$VAR\ (X) = \sigma^2 = \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{n}$$

$$VAR\ (X) = \sigma^2 = \sum_{i=1}^{n} \frac{(x_i)^2}{n} - \mu^2$$

Para datos agrupados: Este caso se presenta cuando tenemos los datos tabulados en las distintas tablas presentadas anteriormente. Usaremos una de las siguientes expresiones

$$VAR\ (X) = \sigma^2 = \sum_{i=1}^{m} \frac{(x_i - \mu)^2 * n_i}{n} \quad m = numero\ de\ filas$$

$$VAR\ (X) = \sigma^2 = \sum_{i=1}^{m} \frac{x_i^2 * n_i}{n} - \mu^2,\ \ m = numero\ de\ filas$$

5.1.3.2. Propiedades de la varianza. La varianza tiene las siguientes propiedades:

- ✓ La varianza de un conjunto de observaciones siempre es un número no negativo.
$$Var(x) \geq 0$$
- ✓ La varianza de una constante es cero.
$$Var(k) = 0$$
- ✓ La varianza de una variable más/menos una constante es igual a la varianza de la variable.
$$Var(x \pm k) = Var(x)$$
- ✓ La varianza de la variable multiplicada por una constante es igual al cuadrado de la constante por la varianza de la variable.
$$Var(kx) = k^2 Var(x)$$

Remarca: La varianza tiene las $[u]^2$

Ejemplo: Determine la varianza de los datos

$$4\$,\quad 6\$,\quad 11\$$$

Solución

$$\mu = \frac{4+6+11}{3} = 7\$$$

$$VAR\ (X) = \sigma^2 = \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{n} = \frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + (x_3 - \mu)^2}{3}$$

$$= \frac{(4-7)^2 + (6-7)^2 + (11-7)^2}{3} = 8.667\ \$^2$$

Ejemplo: Determine la varianza

[L_{i-1} - L_i)	Xi	n_i	h_i (%)	N_i	H_i (%)
[7 a 9.4)	8.2	8	32	8	32
[9.4 a 11.8)	10.6	5	20	13	52
[11.8 a 14.2)	13	4	16	17	68
[14.2 a 16.6)	15.4	6	24	23	92
[16.6 a 19]	17.8	2	8	25	100
TOTAL		25	100		

$$\mu = \frac{8.2 * 8 + 10.6 * 5 + 13 * 4 + 15.4 * 6 + 17.8 * 2}{25} = 11.944$$

$$VAR\ (X) = \sigma^2 = \sum_{i=1}^{m} \frac{(x_i - \mu)^2 * n_i}{n} = \frac{(x_1 - \mu)^2 * n_1 + (x_2 - \mu)^2 * n_2 + \cdots + (x_m - \mu)^2 * n_m}{n}$$

$$= \frac{(8.2-11.944)^2 * 8 + (10.6-11.944)^2 * 5 + (13-11.944)^2 * 4 + (15.4-11.944)^2 * 6 + (17.8-11.944)^2 * 2}{25}$$

$$= 10.635\ [u]^2$$

Desviación Estándar.- En vista de que las unidades de la varianza están elevadas al cuadrado, esto presenta un inconveniente con respecto a la interpretación de esta cantidad. En la práctica otra mediada basada en el valor de la varianza que sirve para dar una medida de dispersión en la misma unidad en la que están los datos. Este indicador es la desviación típica o estándar.

La desviación estándar es la raíz cuadrada positiva de la varianza y esta expresada en las mismas unidades de la variable. Esta medida esta denotada por: $\sigma,\ \sigma_x,\ S$.

Esta medida se calcula mediante la siguiente expresión:

- Para datos no agrupados

$$\sigma = \sqrt{\sum_{i=1}^{n} \frac{(x_i-\mu)^2}{n}}$$

$$\sigma = \sqrt{\sum_{i=1}^{n} \frac{(x_i)^2}{n} - \mu^2}$$

- Para datos agrupados

$$\sigma = \sqrt{\sum_{i=1}^{n} \frac{(x_i-\mu)^2 * n_i}{n}}$$

$$\sigma = \sqrt{\sum_{i=1}^{m} \frac{x_i^2 * n_i}{n} - \mu^2}$$

Ejemplo: Determine la desviación típica o estandar de los datos

$$4\$,\quad 6\$,\quad 11\$$$

Solución

$$\mu = \frac{4+6+11}{3} = 7\$$$

$$VAR\ (X) = \sigma^2 = \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{n} = \frac{(x_1-\mu)^2 + (x_2-\mu)^2 + (x_3-\mu)^2}{3}$$

$$= \frac{(4-7)^2 + (6-7)^2 + (11-7)^2}{3} = 8.667\ \$^2$$

$$\sigma = \sqrt{8.667} = 2.944\ \$$$

Ejemplo: Determine la varianza

[L_{i-1} - L_i)	**Xi**	**n_i**	**h_i (%)**	**N_i**	**H_i (%)**
[7 a 9.4)	8.2	8	32	8	32
[9.4 a 11.8)	10.6	5	20	13	52
[11.8 a 14.2)	13	4	16	17	68
[14.2 a 16.6)	15.4	6	24	23	92
[16.6 a 19]	17.8	2	8	25	100
TOTAL		25	100		

$$\mu = \frac{8.2 * 8 + 10.6 * 5 + 13 * 4 + 15.4 * 6 + 17.8 * 2}{25} = 11.944$$

$$VAR\ (X) = \sigma^2 = \sum_{i=1}^{m} \frac{(x_i - \mu)^2 * n_i}{n} = \frac{(x_1 - \mu)^2 * n_1 + (x_2 - \mu)^2 * n_2 + \cdots + (x_m - \mu)^2 * n_m}{n}$$

$$= \frac{(8.2 - 11.944)^2 * 8 + (10.6 - 11.944)^2 * 5 + (13 - 11.944)^2 * 4 + (15.4 - 11.944)^2 * 6 + (17.8 - 11.944)^2 * 2}{25}$$

$$= 10.635\ [u]^2$$

$$\sigma = \sqrt{10.635} = 3.261\ [u]$$

Medidas de comparación. Para efectuar comparaciones entre dos distribuciones de frecuencia se buscan indicadores que no estén expresados en la escala de la medida de la variable y que solo estén expresados en números.

Estas medidas son llamadas también Medidas de dispersión relativa y las más utilizadas son:

- Coeficiente de variación.
- Variable estandarizada.

Coeficiente de Variación. Es la relación que establece entre la desviación estándar existente en la distribución por unidad de media aritmética.

El coeficiente de variación se calcula mediante la siguiente expresión:

$$Coeficiente\ de\ Variación = C.V. = \frac{\sigma}{\mu}$$

El coeficiente de variación es útil, cuando se compara la variabilidad de dos o más conjuntos de datos que difieren de modo considerable en las magnitudes de las observaciones. Esta fórmula sirve para datos tabulados y no tabulados

Ejemplo: Determine el coeficiente de variación de los datos

$$4\$, \quad 6\$, \quad 11\$$$

Solución

$$\mu = \frac{4+6+11}{3} = 7\$$$

$$VAR\ (X) = \sigma^2 = \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{n} = \frac{(x_1-\mu)^2+(x_2-\mu)^2+(x_3-\mu)^2}{3}$$

$$= \frac{(4-7)^2+(6-7)^2+(11-7)^2}{3} = 8.667\ \$^2$$

$$\sigma = \sqrt{8.667} = 2.944\ \$$$

$$C.V. = \frac{2.944}{7} = 0.421$$

Para su interpretación se lo puede convertir en porcentaje e interpretar

$$C.V. = \frac{2.944}{7} = 0.421\ (42.1\%)\ Variación\ representado\ en\ porcentaje$$

Ejemplo: Determine la varianza

[L_{i-1} - L_i)	**Xi**	**n_i**	**h_i (%)**	**N_i**	**H_i (%)**
[7 a 9.4)	8.2	8	32	8	32
[9.4 a 11.8)	10.6	5	20	13	52
[11.8 a 14.2)	13	4	16	17	68
[14.2 a 16.6)	15.4	6	24	23	92
[16.6 a 19]	17.8	2	8	25	100
TOTAL		25	100		

$$\mu = \frac{8.2*8+10.6*5+13*4+15.4*6+17.8*2}{25} = 11.944$$

$$VAR\ (X) = \sigma^2 = \sum_{i=1}^{m} \frac{(x_i-\mu)^2 * n_i}{n} = \frac{(x_1-\mu)^2 * n_1 + (x_2-\mu)^2 * n_2 + \cdots + (x_m-\mu)^2 * n_m}{n}$$

$$= \frac{(8.2-11.944)^2 * 8 + (10.6-11.944)^2 * 5 + (13-11.944)^2 * 4 + (15.4-11.944)^2 * 6 + (17.8-11.944)^2 * 2}{25}$$

$$= 10.635\ [u]^2$$

$$\sigma = \sqrt{10.635} = 3.261\ [u]$$

$$C.V. = \frac{3.261}{11.944} = 0.273\ (27.3\%)\ \ La\ variación\ representado\ en\ porcentaje$$

Ejemplo: Sean los datos

1, 2, 3

Calcular C.V.

$\mu = 2$

$\sigma = 0.816$

$C.V. = 0.408\ (40.8\%)$

Ejemplo: Sean los datos

501, 502, 503

Calcular C.V.

$\mu = 502$

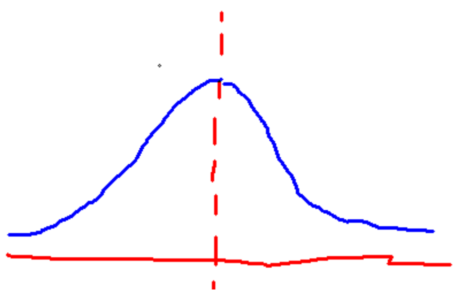
$\sigma = \ 0.816$

$C.V. = 0.00163(0.163\%)$

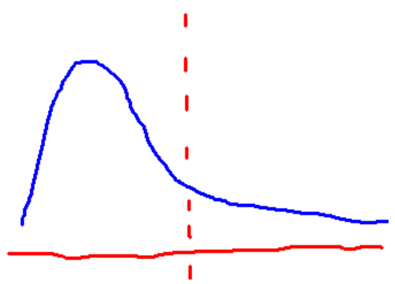
Medidas de Asimetría. En forma general, el grafico de una distribución de frecuencias se relaciona con un eje vertical o un eje de asimetría, el cual atraviesa el punto más alto de la distribución. En términos de este eje de simetría, la representación grafica de una distribución se califica como simetría o asimetría.

El concepto de asimetría de una distribución indica la deformación horizontal de las distribuciones de la frecuencia.

Cuando la grafica de una distribución tiene cola o rama estirada a la derecha o a la izquierda, se dice que es sesgada o asimétrica a la derecha o la izquierda, respectivamente. En este caso la mediana, la media y la moda tienen valores no coincidentes, es decir:



Simétrico



asimetría

Una medida que cuantifica la asimetría, es el momento de tercer orden respecto a la media aritmética, que se obtiene mediante la ecuación:

- Para datos no agrupados:

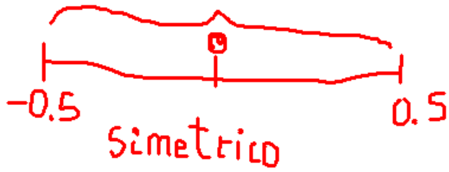
$$tercer\ momento = \ m_3 = \sum_{i=1}^{n} \frac{(x_i - \mu)^3}{n}$$

- Para datos agrupados:

$$tercer\ momento = m_3 = \sum_{i=1}^{n} \frac{(x_i - \mu)^3 * n_i}{n}$$

Si:

Remarca: Puesto que la simetría perfecta es casi de tener, se considerará una variación de 0.5. Siendo 0 la simetría perfecta, considerándose un intervalo cerrado.



Ejemplo: Sean los datos

4, 6, 10, 20

¿Tienen un comportamiento simétrico?

$$\mu = 10$$

$$m_3 = \sum_{i=1}^{n} \frac{(x_i - \mu)^3}{n} = \frac{(4-10)^3 + (6-10)^3 + (10-10)^3 + (20-10)^3}{4} = 180$$

No tienen un comportamiento simétrico